A Roadmap for Grid Performance:

reporting from the frontline

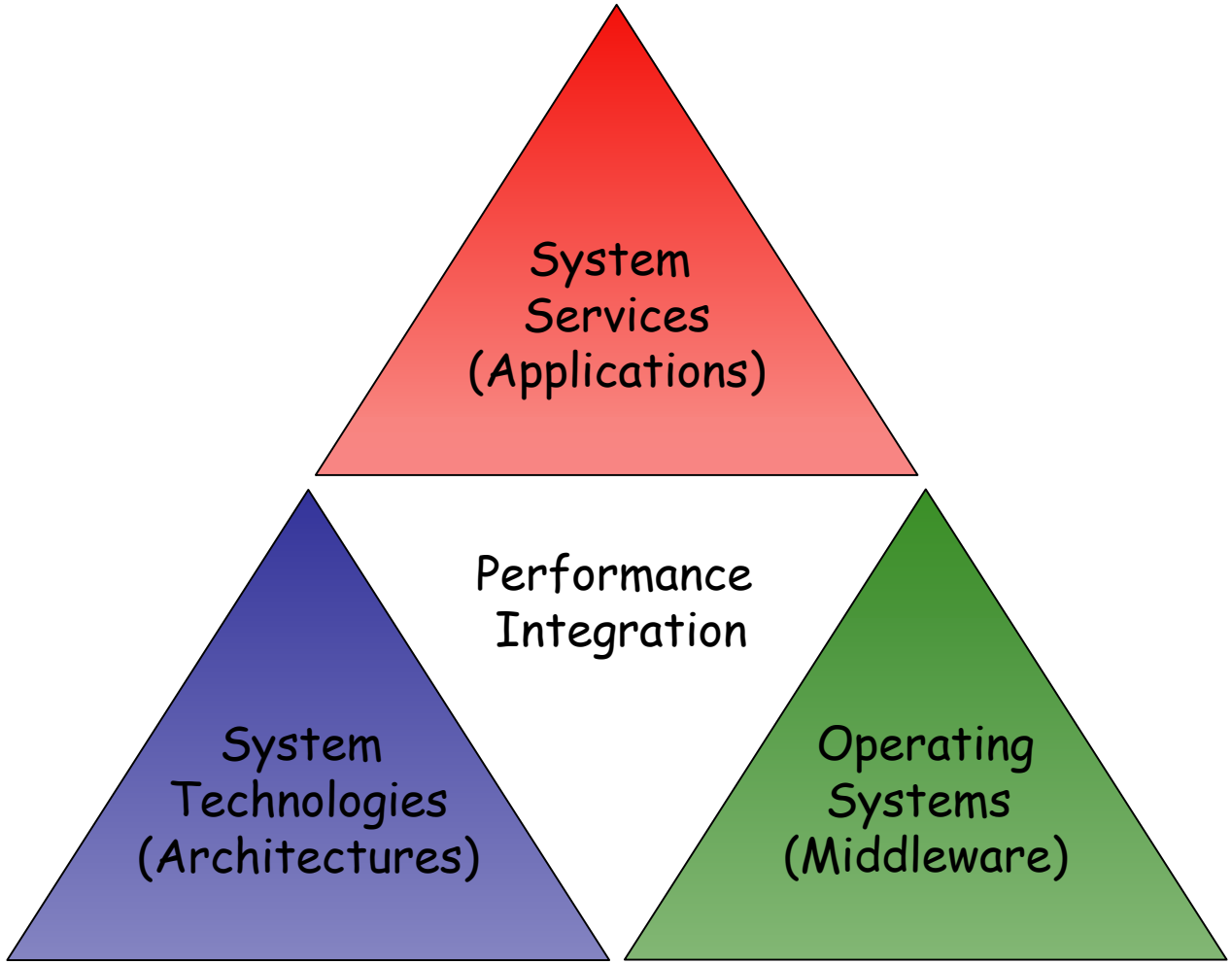
Stephen Jarvis

University of Warwick
Midlands e-Science Centre

Performance-based Middleware for Grid Computing

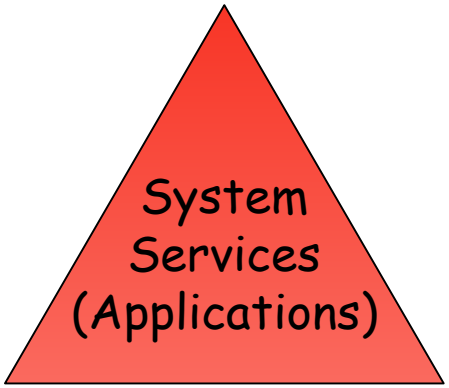
- Integrate established performance tools with emerging grid middleware
  - Understanding requirements of applications
  - Determine what resources are available
  - Matchmaking, while maintaining contract of performance
- Performance-aware services
  - Integrating performance into scheduling, workload management …
  - Developing infrastructure support
  - Benchmarking middleware components
  - Stress-testing these solutions in e-Science framework
- Common theme is *performance awareness*
- Open issues depend on your viewpoint

A Roadmap for Grid Performance



System Services (Applications)

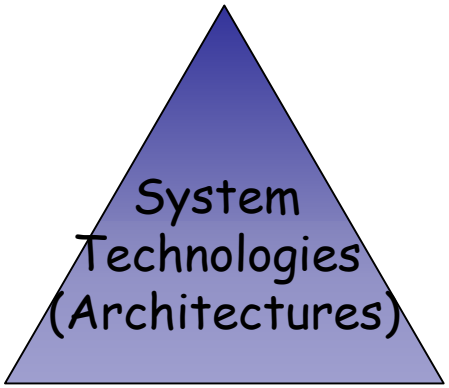
"We will send your result to your burial place…"



- Speed of response
  - Modes of analysis: Monitoring, modelling and simulation
  - Accuracy vs. lifetime vs. overhead
- More general resource consumption
  - Data, networking, computation, power
  - Mapping performance to cost ($)
- Classes of application
  - Scientific vs. business, data vs. compute
  - Performability and QoS
- Design
  - Composition in time (workflow) or space (component assembly)
  - Cost models (flow, sub-flow, tasks,…,flop)
  - Performance-based service selection

System Technologies (Architectures)

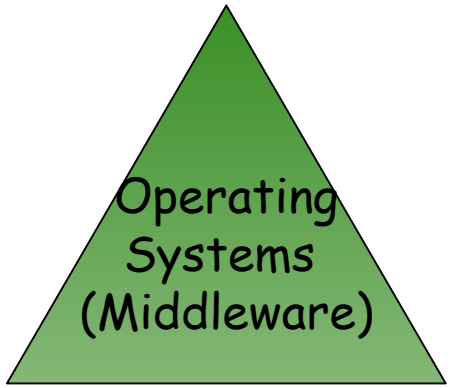
"Prediction is difficult, especially about the future…"



- Representing resources
  - How to represent resources (e.g. power?)
  - Relating this to the application
- Network
  - Bandwidth, response time, TCP perf.
  - Repositories (e.g. GridFTP logs and prediction)
  - Management (provisioning) vs. monitoring
- Data
  - File transfer costs
  - Proximity and staging (JoSH)
- System views
  - Global snapshots
  - Decay vs. quality

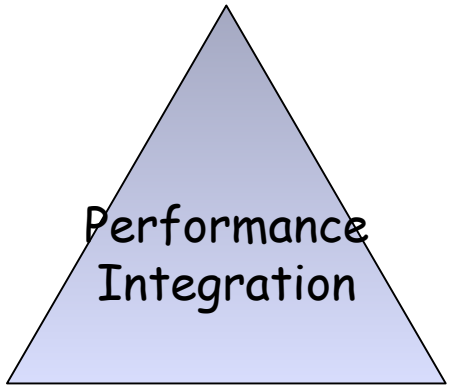
Operating Systems (Middleware)

"Turning off the security features improves performance…"



- Middleware design
  - Additional complexities of the software stack (depth and composition)
  - Benchmarking components (e.g. MDS) and competing implementations
  - Software and architectures evolving
- Middleware activities
  - Scheduling and workload steering
  - Reservation / co-allocation
  - Charging (e.g. GESA-WG)
- Granularity
  - Micro vs. macro performance data
  - Reliability of data and services

Performance Integration



- Information services
  - Metadata definitions
  - Overhead
  - Built on services (e.g. self adapting and optimising perf. monitoring)
  - Capturing perf. differences between VOs
- Logging and book-keeping
  - Logging infrastructure
  - Resource usage profiles (e.g. GridMon)
  - Performance pathologies
- Grid benchmarking (GB-WG)
  - Test-beds and load generators
- Scenario testing
  - Adding / removing resources
  - Adding / removing / updating services

- What differentiates Grid performance research from classical performance research?
  - i.e. what existing work can be applied?, what new work is needed?
  - Techniques can be reused
  - Collaborative process (people, systems) is perhaps new
  - No greenroom, no one solution fits all

- Challenges and Opportunities in Grid Performability
  www.cs.ncl.ac.uk/research/pubs/trs/papers/842.pdf
- e-Science Performance Engineering Workshop
  www.nesc.ac.uk/action/esi/contribution.cfm?Title=127
- Open Issues in Grid Scheduling
  www.nesc.ac.uk/action/esi/contribution.cfm?Title=309
- High Performance Systems Group, Warwick
  www.dcs.warwick.ac.uk/research/hpsg